AFRL-SN-RS-TR-1998-46
Interim Technical Report
March 1998

ADAPTIVE SIDELOBE CANCELING USING COMPLEX-VALUED CANONICAL VARIABLES

Scientific Studies Corp.

Jaime R. Roman

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AIR FORCE RESEARCH LABORATORY
SENSORS DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

AFRL-SN-RS-TR-1998-46 has been reviewed and is approved for publication.

APPROVED: James H. Michels
JAMES H. MICHELS
Project Engineer

FOR THE DIRECTOR: Robert G. Polce
ROBERT G. POLCE, Acting Chief
Rome Operations Office
Sensors Directorate

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 1998	Interim Apr 97 - Dec 97

4. TITLE AND SUBTITLE

ADAPTIVE SIDELOBE CANCELING USING COMPLEX-VALUED CANONICAL VARIABLES

5. FUNDING NUMBERS

C - F30602-97-C-0085
PE - 65502F
PR - 3005
TA - 72
WU - 16

6. AUTHOR(S)

Jaime R. Roman

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Scientific Studies Corporation
2250 Quail Ridge
Palm Beach Gardens FL 33418

8. PERFORMING ORGANIZATION REPORT NUMBER

SSC-TR-98-01

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFRL/SNRT
26 Electronic Pky
Rome NY 13441-4514

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

AFRL-SN-RS-TR-1998-46

11. SUPPLEMENTARY NOTES

AFRL Project Engineer: James H. Michels/SNRT/(315) 330-4432

12a. DISTRIBUTION AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Report developed under SBIR contract. A sidelobe canceler is an adaptive array configuration structured to enhance the reception of a desired signal corrupted by unwanted interference signals in applications such as radar, sonar, and communication systems. Several performance criteria, including the minimum mean-square error (MMSE) and maximum likelihood, have been defined for adaptive array problems. A new performance criterion, referred to as maximum correlation (MC), is introduced herein in the context of sidelobe canceling. The MC criterion is based on the concept of canonical variables introduced by Hotelling in 1936 to define a canonical relationship between two sets (or vectors) of real-valued random variables. In this report Hotelling's formulation is extended to handle complex-valued random variables for the special case of a complex-valued scalar and a complex-valued vector. In addition, it is shown that the weight vector solution for the MC criterion is equivalent, up to a scale factor, to the MMSE weight vector.

14. SUBJECT TERMS	15. NUMBER OF PAGES
SBIR Report; adaptive arrays; sidelobe canceling; Canonical variables; Canonical correlations	36
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

TABLE OF CONTENTS

LIST OF FIGURES

1.0 INTRODUCTION

This documents is an interim report for Phase II of the "Two-Dimensional Processing for Radar Systems" Small Business Innovation Research (SBIR) program for the Radar Signal Processing Branch of the Sensors Directorate of the Air Force Research Laboratory (AFRL). The first version of the material presented herein appeared as two internal and proprietary technical memoranda at Scientific Studies Corporation (SSC) several years ago (Román, 1991; 1992). This report is a revised version of that earlier work at SSC, due to its relevance to the aforementioned SBIR program.

Adaptive arrays provide a means for enhancing the reception of a desired signal in the presence of unwanted interference signals in applications such as radar, sonar, and communications systems. A specific configuration for such a system is the sidelobe canceler, as illustrated in Figure 1-1. In a sidelobe canceler a narrow-beam, low-sidelobe main sensor is oriented to receive a desired signal from a known direction, and multiple auxiliary sensors receive the interference signals that enter the main sensor through its sidelobes. In the specific configuration considered here, each auxiliary sensor has a single null which is oriented in the direction of the desired signal. Thus, the main sensor receives all signals (desired as well as interference), but the auxiliary sensors receive only the interference signals. Appropriate selection of the weight values in a weighted linear combination of all the sensors results in combined-pattern nulls in each of the directions of the interference sources, and thus allows for enhanced desired signal reception and unwanted signal rejection.

Several performance criteria have been defined in the literature for adaptive array problems, along with various

adaptive algorithms for the computation of the weights. The most common performance criteria are the following: a) minimum mean-square error (MMSE) criterion (Wiener filter), b) signal-to-noise ratio (SNR) criterion, c) maximum likelihood (ML) criterion, and d) minimum variance (MV) criterion. Each of these leads to an optimum steady-state solution for the adaptive weights, and under appropriate conditions all the solutions are equivalent.

A new performance criterion, referred to as maximum correlation (MC), is introduced herein. The MC criterion is based on the concept of canonical correlations and canonical variates, which was introduced by Hotelling (1936) to define a canonical relationship between two sets (or vectors) of real-valued random variables (see also Anderson [1958]). In Hotelling's formulation the canonical variables embody the essence of the correlation structure among the variables of the two given sets. The approach presented herein parallels the work of Hotelling (1936), but also provides an important extension to handle complex-valued random variables. Specifically, the canonical variables technique is established for the special case of a complex-valued scalar and a complex-valued vector. Also, it is shown that the optimum solution for the MC criterion is equivalent, up to a scalar factor, to the MMSE (Wiener filter) solution.

Hotelling (1936) defined the canonical correlations technique for the case of two real-valued vectors. The derivation presented herein is for the case of a complex-valued scalar and a complex-valued vector, but it can be extended to the case of two complex-valued vectors of arbitrary dimensions. The general case will be considered in a separate report.

1.1 <u>Notation</u>

Mathematical symbols for variables, constants, operators, etc., are represented in either Helvetica font, Times font, or

Symbol font. An underscored lower-case variable, $\underline{x}$, denotes a vector, and the transpose operator is denoted by a superscript upper-case T. The complex conjugate operator is denoted by an overbar, $\bar{x}$, and the complex conjugate transpose operator is denoted by a superscript dagger in Times font, †. Vertical bars, $|\bullet|$, denote the determinant of the matrix enclosed within the bars, or the absolute value of the scalar enclosed within the bars. The expected value operator is denoted as $E[\bullet]$. O_N denotes an NxN null matrix, and I_N denotes the NxN identity matrix. The real and imaginary components of a complex-valued quantity are denoted as $\mathrm{Re}\{\bullet\}$ and $\mathrm{Im}\{\bullet\}$, respectively.

1.2 Report Overview

Sidelobe cancelation using canonical correlations is defined in Section 2.0, following the standard function minimization approach. The MC solution to the sidelobe canceling problem is derived in Section 3.0 using the complex gradient operator discussed by Spiegel (1964). And the MMSE solution is derived analogously in Section 4.0. Two variations of the MC solution (obtained via modifications to the performance index utilized in Section 3.0) are derived in Section 5.0. The main results are summarized and comments are presented in Section 6.0. An analytical derivation of the performance index adopted in Section 2.0 is presented in Appendix A.

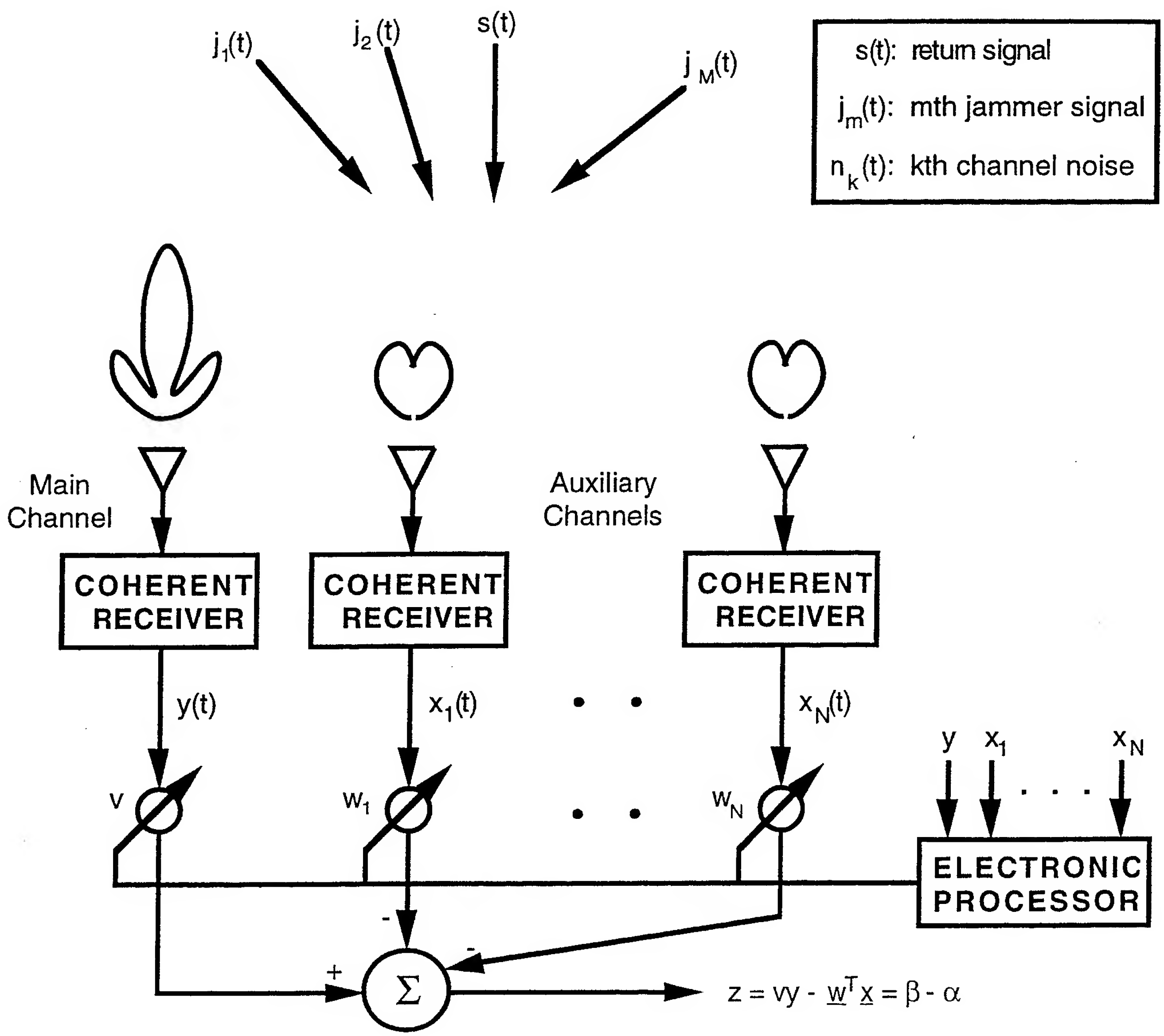

Figure 1-1. Sidelobe canceler configuration for maximum correlation (MC) approach.

2.0 PROBLEM FORMULATION

The cases considered herein involve spatially-distinct narrow-band signals (desired as well as interference), which allows the use of the complex representation for the signal in each receiver channel. Thus, referring to Figure 1-1, $y(t)$ is a complex-valued random process and $\underline{x}(t)$ is a complex-valued vector random process. Each receiver channel adds sensor noise to the received signal, and it is assumed that the sensor noise is independent from channel to channel. Further, the sensor noise is independent from the desired signal as well as from the interference signals. For simplicity, $y(t)$ and $\underline{x}(t)$ are assumed to have mean zero, and all processes are assumed to be stationary in the wide-sense (temporal variations are handled by appropriately restricting the duration of the processing interval). The weights $v(t)$ and $\underline{w}(t)$ are complex-valued quantities also. It is assumed that the desired signal is independent of the interference signals. A digital processor is assumed, so the input to the processor consists of the received signals sampled at equi-spaced discrete instants of time, and can be represented as

$$y(kT) = s(kT) + r_0(kT) + n_0(kT) \tag{2-1}$$

$$\underline{x}(kT) = \underline{r}(kT) + \underline{n}(kT) \tag{2-2}$$

where T denotes the sampling interval and k is an integer. Also, $r_0(kT)$ denotes the combination of all the interference signals arriving at the main channel, $n_0(kT)$ denotes the channel noise in the main channel, $\underline{r}(kT)$ represents the combination of all the interference signals arriving at the auxiliary channels, and $\underline{n}(kT)$ represents the channel noise in the auxiliary channels. For notational simplicity, the explicit dependence on time will be dropped for all variables, and the variables in equations will be understood to taken at the same sampling instant.

In the system configuration shown in Figure 1-1 the array output signal

$$z = vy - \underline{w}^T\underline{x} = \beta - \alpha \tag{2-3}$$

is an accurate estimate of the desired signal component present in y if the weights v and $\underline{w}$ are selected such that an appropriate criterion is optimized. As stated earlier, several criteria have been formulated and presented in the literature (MMSE; SNR; ML; MV). A new, alternative criterion is formulated next.

The MC criterion is based on the postulate that the array output signal, z, is also an accurate estimate of the desired signal component present in y if the weights v and $\underline{w}$ are selected such that the correlation between $\alpha = \underline{w}^T\underline{x}$ and $\beta = vy$ is maximized, a consequence of the fact that any correlation between y and $\underline{x}$ is due only to the interference signal components present in each. This required maximization of the correlation provides the name for the technique.

In analytical terms, the correlation coefficient between the zero-mean variables α and β is

$$\rho_{\alpha\beta} = \frac{E[\overline{\alpha}\beta]}{E[\overline{\alpha}\alpha]\,E[\overline{\beta}\beta]} = \frac{E[\underline{w}^\dagger\overline{\underline{x}}yv]}{(E[\underline{w}^\dagger\overline{\underline{x}}\underline{x}^T\underline{w}])(E[\overline{v}\overline{y}yv])} = \frac{\underline{w}^\dagger E[\overline{\underline{x}}y]v}{(\underline{w}^\dagger E[\overline{\underline{x}}\underline{x}^T]\underline{w})(\overline{v}E[\overline{y}y]v)} \tag{2-4a}$$

$$\rho_{\alpha\beta} = \frac{\underline{w}^\dagger\underline{a}v}{(\underline{w}^\dagger C\underline{w})(\overline{v}bv)} \tag{2-4b}$$

where the complex-valued vector $\underline{a}$ denotes the covariance between $\underline{x}$ and y, the real-valued scalar b denotes the covariance of y, and the Hermitian matrix C denotes the covariance of $\underline{x}$. The cross-covariance $\underline{a}$ is the cross-covariance of the interference signals,

(2-5) $\underline{a} = E[\underline{\bar{x}}y] = E[(\underline{\bar{r}} + \underline{\bar{n}})(s + r_0 + n_0)] = E[\underline{\bar{r}}\, r_0]$

In turn, covariances b and C are determined as

(2-6) $b = E[\bar{y}y] = E[(\bar{s} + \bar{r}_0 + \bar{n}_0)(s + r_0 + n_0)] = E[\bar{s}s] + E[\bar{r}_0 r_0] + E[\bar{n}_0 n_0]$

(2-7) $C = E[\underline{\bar{x}}\underline{x}^T] = E[(\underline{\bar{r}} + \underline{\bar{n}})(\underline{r}^T + \underline{n}^T)] = E[\underline{\bar{r}}\,\underline{r}^T] + E[\underline{\bar{n}}\underline{n}^T]$

Since the correlation coefficient of a multiple of α and a multiple of β is the same as the correlation coefficient of α and β, then v and $\underline{w}$ can be selected such that α and β have unit variance (α and β are zero-mean variables). Thus, the correlation coefficient $\rho_{\alpha\beta}$ is to be maximized subject to the constraints that

(2-8) $E[\bar{\alpha}\alpha] = \underline{w}^\dagger E[\underline{\bar{x}}\underline{x}^T]\underline{w} = \underline{w}^\dagger C\underline{w} = 1$

(2-9) $E[\bar{\beta}\beta] = \bar{v}E[\bar{y}y]v = \bar{v}bv = 1$

where b and C are as defined above.

The correlation coefficient between α and β is a complex number, and its maximum possible magnitude is equal to 1. For any normalized complex number, maximum magnitude can be attained with an infinity of arguments (the angle part of the complex number). Furthermore, maximization of $\rho_{\alpha\beta} + \bar{\rho}_{\alpha\beta}$ is equivalent to maximization of $\rho_{\alpha\beta}$ for a particular choice of angle in the complex plane, as suggested in Figure 2-1. Therefore, it is reasonable to state the sidelobe canceler problem as follows: optimize the performance criterion

(2-10) $J(v,\underline{w}) = \frac{1}{2}(\rho_{\alpha\beta} + \bar{\rho}_{\alpha\beta}) = \frac{1}{2}(\underline{w}^\dagger \underline{a}v + \underline{w}^T\underline{\bar{a}}\bar{v}) = \mathrm{Re}\{\underline{w}^\dagger \underline{a}v\}$

subject to the following two constraints:

(2-11) $L_1(v) = \bar{v}bv - 1 = 0$

(2-12) $L_2(\underline{w}) = \underline{w}^{\dagger}C\underline{w} - 1 = 0$

Combination of the performance criterion (2-10) with the above constraints leads to the following real-valued functional which is to be optimized with respect to the complex variables v and $\underline{w}$:

(2-13) $$H(v,\underline{w}) = J(v,\underline{w}) - \frac{1}{2}\lambda_1 L_1(v) - \frac{1}{2}\lambda_2 L_2(\underline{w})$$

(2-14) $$H(v,\underline{w}) = \frac{1}{2}(\underline{w}^{\dagger}\underline{a}v + \underline{w}^{T}\overline{\underline{a}v}) - \frac{1}{2}\lambda_1(\bar{v}bv - 1) - \frac{1}{2}\lambda_2(\underline{w}^{\dagger}C\underline{w} - 1)$$

where λ_1 and λ_2 are real-valued Lagrange multipliers. The scalar functional $H(v,\underline{w})$ is referred to as the Hamiltonian.

In optimization problems involving complex-valued variables it is important to define the Hamiltonian function (performance index and constraints) as a real-valued function of the complex-valued variables. This allows for a unique value for the optimum performance index in well-defined problems (in problems with a complex-valued criterion the maximum or minimum may not be unique; for example, the complex numbers $1 - j2$ and $2 + j1$ have the same magnitude but are distinct values). A real-valued Hamiltonian also allows for straightforward application of complex gradient operators such as those considered herein as well as in most of the literature.

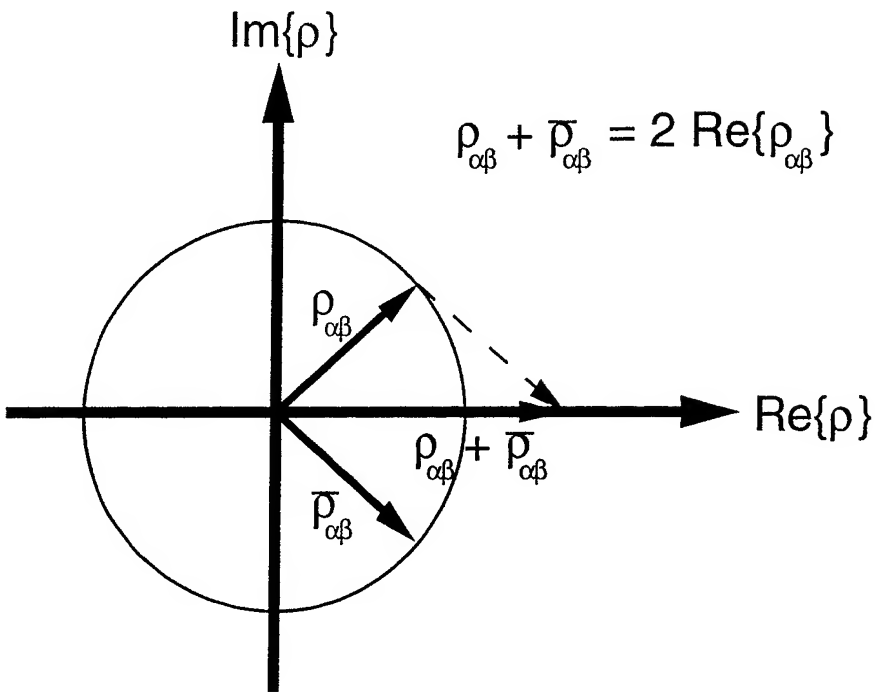

Figure 2-1. Complex correlation coefficient and optimization criterion.

3.0 SIDELOBE CANCELER MAXIMUM CORRELATION SOLUTION

Optimization of the functional $H(v,\underline{w})$ is carried out by taking the gradient of $H(v,\underline{w})$ with respect to v and to $\underline{w}$, and setting these gradients equal to zero. The resulting equations, together with the constraint equations, are solved for the values of v and $\underline{w}$ that optimize the performance index. An optimal solution is a maximum or minimum depending on the sign of the second derivative of the Hamiltonian at that solution.

Since $H(v,\underline{w})$ is a real-valued function of the complex-valued variables v and $\underline{w}$, its gradient with respect to each of these variables is a complex-valued quantity. In this report the following complex gradient operator is used (Spiegel, 1964):

(3-1) $$\nabla_v H(v,\underline{w}) = \frac{\partial H(v,\underline{w})}{\partial v_r} + j\frac{\partial H(v,\underline{w})}{\partial v_i}$$

(3-2) $$\nabla_{\underline{w}} H(v,\underline{w}) = \frac{\partial H(v,\underline{w})}{\partial \underline{w}_r} + j\frac{\partial H(v,\underline{w})}{\partial \underline{w}_i}$$

where the subscripts r and i denote the real and imaginary parts, respectively, of v or $\underline{w}$. Calculation of these gradients and setting them equal to zero gives two equations which, together with the two constraint equations, constitute the necessary conditions which must be satisfied by the optimum values of v and $\underline{w}$. Specifically,

(3-3) $$\nabla_v H(v,\underline{w}) = \underline{a}^{\dagger}\underline{w} - \lambda_1 bv = 0$$

(3-4) $$\nabla_{\underline{w}} H(v,\underline{w}) = \underline{a}v - \lambda_2 C\underline{w} = \underline{0}$$

(3-5) $$\nabla_{\lambda_1} H(v,\underline{w}) = \frac{1}{2}(\overline{v}bv - 1) = 0$$

(3-6) $$\nabla_{\lambda_2} H(v,\underline{w}) = \frac{1}{2}(\underline{w}^{\dagger} C \underline{w} - 1) = 0$$

Multiplication of Equation (3-3) by $\bar{v}$ and using Equation (3-5) allows λ_1 to be obtained as:

(3-7) $$\lambda_1 = \bar{v}\underline{a}^{\dagger}\underline{w} = \underline{w}^{T}\overline{\underline{a}v}$$

Similarly, Equation (3-4) is multiplied from the left by $\underline{w}^{\dagger}$, and together with Equation (3-6) is used to solve for λ_2 as:

(3-8) $$\lambda_2 = \underline{w}^{\dagger}\underline{a}v$$

Notice that $\overline{\lambda_1} = \lambda_2$, and recall that both of these multipliers are real-valued since the Hamiltonian in Equation (2-13) is real-valued; therefore, it follows that

(3-9) $$\lambda_2 = \lambda_1 = \lambda = \mathrm{Re}\{\underline{w}^{\dagger}\underline{a}v\}$$

Furthermore, it is also true that

(3-10) $$\mathrm{Im}\{\underline{w}^{\dagger}\underline{a}v\} = 0$$

Equation (3-10) is a reasonable result given the selected performance index. Notice that relation (3-9) implies the following equality (recall Equation (2-10)):

(3-11) $$\lambda = J(v,\underline{w})$$

This last result is interesting because it indicates that the Lagrange multipliers are equal to the performance criteria to be optimized. An identical result is true in canonical correlations for real-valued random variates (Anderson, 1958).

Even though the value of λ is unknown (since v and $\underline{w}$ are undetermined), Equation (3-9) allows simplification of the gradient equations, as shown next. Substitution of λ for λ_1 and λ_2 into the necessary conditions (3-3) and (3-4) leads to the following partitioned matrix relation:

$$\text{(3-12)} \qquad \begin{bmatrix} -\lambda b & \underline{a}^{\dagger} \\ \underline{a} & -\lambda C \end{bmatrix} \begin{bmatrix} v \\ \underline{w} \end{bmatrix} = \underline{0}$$

Equation (3-12) represents a generalized eigenvector/eigenvalue problem,

$$\text{(3-13)} \qquad A\underline{e} = \lambda B\underline{e}$$

wherein the following partitioned quantities have been defined implicitly:

$$\text{(3-14)} \qquad A = \begin{bmatrix} 0 & \underline{a}^{\dagger} \\ \underline{a} & O_N \end{bmatrix}$$

$$\text{(3-15)} \qquad B = \begin{bmatrix} b & \underline{0}^T \\ \underline{0} & C \end{bmatrix}$$

$$\text{(3-16)} \qquad \underline{e} = \begin{bmatrix} v \\ \underline{w} \end{bmatrix}$$

The presence of matrix B in equation (3-13) is the reason for the "generalized" terminology. This problem can be converted to a conventional eigenvector/eigenvalue problem because matrix B is non-singular ($b \neq 0$ and matrix C is non-singular due to the sensor noise present in each channel). Specifically,

$$\text{(3-17)} \qquad B^{-1}A\underline{e} = \lambda\underline{e}$$

or equivalently,

$$[\lambda I_{N+1} - B^{-1}A]\,\underline{e} = \underline{0} \tag{3-18}$$

The matrix product $B^{-1}A$ is determined trivially as

$$B^{-1}A = \begin{bmatrix} 0 & \left(\frac{1}{b}\right)\underline{a}^{\dagger} \\ C^{-1}\underline{a} & O_N \end{bmatrix} \tag{3-19}$$

The eigenvalues of $B^{-1}A$ are the solutions to the characteristic equation associated with the determinant of the matrix in Equation (3-19); that is, the eigenvalues of $B^{-1}A$ are the roots of the following equation:

$$\left|\lambda I_{N+1} - B^{-1}A\right| = \begin{vmatrix} \lambda & -\left(\frac{1}{b}\right)\underline{a}^{\dagger} \\ -C^{-1}\underline{a} & \lambda I_N \end{vmatrix} = 0 \tag{3-20}$$

Examination of the structure of the $(N+1)x(N+1)$ matrix $B^{-1}A$ indicates that its rank is two for all cases (only two linearly-independent rows or columns). Thus, $N-1$ of its $N+1$ eigenvalues are zero. All the eigenvalues can be obtained via the application of the generalized algorithm of Gauss for partitioned matrices (Gantmacher, 1977), which leads to the following equation:

$$\left|\lambda I_{N+1} - B^{-1}A\right| = \left(\lambda^{N-1}\right)\left(\lambda^2 - \frac{\underline{a}^{\dagger}C^{-1}\underline{a}}{b}\right) = 0 \tag{3-21}$$

Equation (3-21) clearly indicates that $N-1$ eigenvalues are equal to zero, and that two real-valued, non-zero eigenvalues are given by

$$\lambda = \pm\sqrt{\frac{\underline{a}^{\dagger}C^{-1}\underline{a}}{b}} \tag{3-22}$$

Of these two values, the positive one is selected since the maximum value is sought for the performance criterion (recall that

$J(v,\underline{w}) = \lambda$). The eigenvector corresponding to the selected eigenvalue is obtained by substituting the selected value for λ into the eigenvalue/eigenvector relation, Equation (3-17), which results in the following two equations:

$$\left(\frac{1}{b}\right)\underline{a}^{\dagger}\underline{w} = \sqrt{\frac{\underline{a}^{\dagger}C^{-1}\underline{a}}{b}}\; v \tag{3-23}$$

$$C^{-1}\underline{a}v = \sqrt{\frac{\underline{a}^{\dagger}C^{-1}\underline{a}}{b}}\; \underline{w} \tag{3-24}$$

Simple algebraic manipulations indicate that these two equations are equivalent. This indicates that any one of the unknown parameters (v or any element of $\underline{w}$) can be selected arbitrarily, which is the case always in eigenvalue/eigenvector problems. Using Equation (3-24) to solve for $\underline{w}$ results in

$$\underline{w} = \sqrt{\frac{b}{\underline{a}^{\dagger}C^{-1}\underline{a}}}\; C^{-1}\underline{a}v \tag{3-25}$$

which indicates that the weight vector $\underline{w}$ is a linear function of the scalar weight v. The selection

$$v_{MC} = \frac{1}{\sqrt{b}} \tag{3-26}$$

satisfies the constraint for v (namely, $\bar{v}vb = 1$), and leads to a simple expression for the MC weight vector $\underline{w}_{MC}$:

$$\underline{w}_{MC} = \frac{1}{\sqrt{\underline{a}^{\dagger}C^{-1}\underline{a}}}\, C^{-1}\underline{a} \tag{3-27}$$

The maximum correlation value is determined given the expressions for the weights, and the result is

$$\rho_{\alpha\beta} = \underline{w}_{MC}^{\dagger}\underline{a}v_{MC} = \sqrt{\frac{\underline{a}^{\dagger}C^{-1}\underline{a}}{b}} \tag{3-28}$$

Notice that the correlation is a real-valued quantity, as expected (recall Equation (3-10)). Furthermore, Equations (2-10), (3-9), (3-22), and (3-28) imply that

(3-29) $$\rho_{\alpha\beta} = J(v_{MC}, \underline{w}_{MC}) = \lambda$$

A relation identical to (3-29) is true also in the real-valued case of canonical variates and canonical correlations (Anderson, 1958).

4.0 SIDELOBE CANCELER MINIMUM MEAN-SQUARE ERROR SOLUTION

Consider now the sidelobe canceler problem formulated using the MMSE criterion. For this case the main channel weight is set equal to unity; that is, $v = 1$. The performance criterion to be optimized is the following real-valued functional,

$$J_0(\underline{w}) = E\left[\,|z|^2\,\right] = E[\bar{z}z] = E\left[\left(\bar{y} - \underline{w}^\dagger \underline{\bar{x}}\right)\left(y - \underline{x}^T \underline{w}\right)\right] \tag{4-1}$$

Functional $J_0(\underline{w})$ can be expressed as (after expanding into product terms, applying the expected value operator, and using the definitions for the covariances b and C, and the cross-covariance $\underline{a}$)

$$J_0(\underline{w}) = b - \underline{w}^\dagger \underline{a} - \underline{a}^\dagger \underline{w} + \underline{w}^\dagger C \underline{w} \tag{4-2}$$

Constraints are not introduced in the conventional MMSE formulation, so the optimization is applied directly to the performance criterion $J_0(\underline{w})$.

As in the MC case, the MMSE performance criterion is a real-valued function of the vector complex-valued variable $\underline{w}$; thus, the complex gradient operator and the optimization approach defined in Section 3.0 apply. Determination of the complex gradient of $J_0(\underline{w})$ with respect to $\underline{w}$ leads to the following necessary condition for an extremum:

$$\nabla_{\underline{w}} J_0(\underline{w}) = -\,2\,\underline{a} + 2\,C\underline{w} = \underline{0} \tag{4-3}$$

This condition is stated equivalently as

$$C\underline{w} = \underline{a} \tag{4-4}$$

Since matrix C is non-singular, the MMSE solution is obtained

directly as

$$\underline{w}_{MSE} = C^{-1}\underline{a} \tag{4-5}$$

Comparison of the MMSE optimal weights with the MC optimal weights shows that both solutions are the same except for a scalar factor present in the the MC solution, Equation (3-27). The optimum performance index is determined by substituting Equation (4-5) into Equation (4-2), which results in

$$J_0(\underline{w}_{MSE}) = b - \underline{a}^{\dagger}C^{-1}\underline{a} \tag{4-6}$$

This expression is quite different from the expression for the MC performance index,

$$J(v_{MC},\underline{w}_{MC}) = \sqrt{\frac{\underline{a}^{\dagger}C^{-1}\underline{a}}{b}} \tag{4-7}$$

5.0 ALTERNATIVE MAXIMUM CORRELATION PERFORMANCE INDICES

The MC optimization problem can be formulated using alternative performance criteria, and the procedure leads to comparable results. In particular, two other such criteria are considered briefly herein.

5.1 Imaginary Part of Correlation Coefficient

The first alternative performance criterion to be optimized is defined as the imaginary part of the correlation coefficient $\rho_{\alpha\beta}$; that is,

$$(5\text{-}1) \qquad J_1(v,\underline{w}) = \frac{1}{2}(\rho_{\alpha\beta} - \bar{\rho}_{\alpha\beta}) = \frac{1}{2}(\underline{w}^\dagger \underline{a}v - \underline{w}^T \overline{\underline{a}v}) = \mathrm{Im}\{\underline{w}^\dagger \underline{a}v\}$$

Substitution of this real-valued criterion in the Hamiltonian (2-13) and determining the optimum solution leads to results similar, but not identical, to those of Section 3.0. Specifically, the real-valued Lagrange multipliers are given as

$$(5\text{-}2) \qquad \lambda_1 = \lambda_2 = \lambda = \mathrm{Im}\{\underline{w}^\dagger \underline{a}v\} = J_1(v,\underline{w})$$

$$(5\text{-}3) \qquad \lambda = \sqrt{\frac{\underline{a}^\dagger C^{-1}\underline{a}}{b}}$$

and the optimum weights are related according to

$$(5\text{-}4) \qquad \underline{w} = -j\sqrt{\frac{b}{\underline{a}^\dagger C^{-1}\underline{a}}}\; C^{-1}\underline{a}v$$

Equation (5-2) is analogous to Equations (3-9) and (3-11), and Equation (5-3) is identical to Equation (3-22). However, Equation (5-4) differs from Equation (3-25) by the $-j$ factor. As before, the scalar weight v can be selected arbitrarily, subject to the normalization constraint. A reasonable selection is

(5-5) $$v_{MC} = j \frac{1}{\sqrt{b}}$$

which satisfies the normalization constraint ($\bar{v}vb = 1$) and results in the same expression for the weight vector $\underline{w}_{MC}$ as in Equation (3-27),

(5-6) $$\underline{w}_{MC} = \frac{1}{\sqrt{\underline{a}^{\dagger}C^{-1}\underline{a}}} C^{-1}\underline{a}$$

The maximum correlation value is determined given the MC expressions for the weights, and results in the value

(5-7) $$\rho_{\alpha\beta} = \underline{w}^{\dagger}_{MC}\underline{a}v_{MC} = j \sqrt{\frac{\underline{a}^{\dagger}C^{-1}\underline{a}}{b}}$$

This correlation is an imaginary-valued quantity, in contrast to the case in Section 3.0. Also, from Equations (5-2), (5-3), and (5-7) it follows that

(5-8) $$\rho_{\alpha\beta} = j\, J_1(v_{MC},\underline{w}_{MC}) = j\, \lambda$$

This is in contrast to Equation (3-29), and is not analogous to the real-valued case of canonical variates and canonical correlations (Anderson, 1958).

5.2 Real Plus Imaginary Parts of Correlation Coefficient

Consider now a second alternative performance criterion to be optimized. In this case the criterion is defined as the sum of the criteria for the two MC cases considered before; that is,

(5-9) $$J_2(v,\underline{w}) = \frac{1}{2}(\rho_{\alpha\beta} + \bar{\rho}_{\alpha\beta}) + \frac{1}{2}(\rho_{\alpha\beta} - \bar{\rho}_{\alpha\beta}) = \mathrm{Re}\{\underline{w}^{\dagger}\underline{a}v\} + \mathrm{Im}\{\underline{w}^{\dagger}\underline{a}v\}$$

As before, substitution of this real-valued criterion in the Hamiltonian (2-13) and determining the optimum solution leads to results similar, but not identical, to those of Section 3.0. Specifically, the real-valued Lagrange multipliers are obtained as

(5-10) $$\lambda_1 = \lambda_2 = \lambda = \mathrm{Re}\{(1 - j)\, \underline{w}^\dagger \underline{a} v\} = \mathrm{Re}\{\underline{w}^\dagger \underline{a} v\} + \mathrm{Im}\{\underline{w}^\dagger \underline{a} v\} = J_2(v, \underline{w})$$

(5-11) $$\lambda = \sqrt{2 \frac{\underline{a}^\dagger C^{-1} \underline{a}}{b}}$$

Notice the $\sqrt{2}$ factor in λ, which is not present in the other two cases, and notice that Equation (5-10) is analogous to Equations (3-9) and (3-11). Also, in this case the MC weights are related according to

(5-12) $$\underline{w} = \frac{(1 - j)}{\sqrt{2}} \sqrt{\frac{b}{\underline{a}^\dagger C^{-1} \underline{a}}}\; C^{-1} \underline{a} v$$

This relation differs from Equation (3-25) by the $1 - j$ factor in the numerator and the $\sqrt{2}$ factor in the denominator. As before, the weight v can be selected arbitrarily, subject to the normalization constraint. A selection compatible with the prior selections is

(5-13) $$v = \frac{(1 + j)}{\sqrt{2b}}$$

which satisfies the normalization constraint ($\bar{v} v b = 1$) and results in the same expression for the weight vector $\underline{w}_{MC}$ as in Equation (3-27); namely,

(5-14) $$\underline{w}_{MC} = \frac{1}{\sqrt{\underline{a}^\dagger C^{-1} \underline{a}}} C^{-1} \underline{a}$$

The maximum correlation value is determined given the MC expressions for the weights, and results in the value

(5-15) $$\rho_{\alpha\beta} = \underline{w}^{\dagger}_{MC}\underline{a}v_{MC} = \frac{(1+j)}{\sqrt{2}}\sqrt{\frac{\underline{a}^{\dagger}C^{-1}\underline{a}}{b}} = \frac{(1+j)}{2}\, J_2(v_{MC},\underline{w}_{MC}) = \frac{(1+j)}{2}\,\lambda$$

which is a complex-valued quantity. Notice that in this case $\mathrm{Re}\{\underline{w}^{\dagger}\underline{a}v\} = \mathrm{Im}\{\underline{w}^{\dagger}\underline{a}v\}$, and as a result, the optimized value of the performance index is

(5-16) $$J_2(v_{MC},\underline{w}_{MC}) = 2\, J(v_{MC},\underline{w}_{MC}) .$$

Clearly, the approach based on the original MC performance index (Section 3.0) is preferred.

6.0 SUMMARY AND COMMENTS

A new performance criterion has been presented for the adaptive sidelobe canceler problem. The criterion is based on the canonical correlations and canonical variates concept developed by Hotelling (1936), and thus it is referred to as the maximum correlation (MC) criterion. The optimal MC weights are derived, and it is shown that the MC solution is equivalent to the MMSE solution. Alternative MC solutions based on distinct performance indices are presented and compared with the preferred MC solution. The comparison indicates that the selected MC solution (Section 3.0) best matches the characteristics of the real-valued canonical correlations approach as defined by Hotelling (1936).

The complex gradient operator defined by Spiegel (1964) is applied in the derivation of the MC weights. Brandwood (1983) has defined a complex gradient operator for real-valued functions of complex variables different from the one adopted in this report. The Brandwood complex gradient operator is based on a new definition of partial derivative for complex variables, and has some characteristics similar to those of the real gradient operator. Application of the Brandwood complex gradient operator (the "conjugate" form of the operator, denoted as $\nabla_{\overline{V}}(\bullet)$ and $\nabla_{\overline{W}}(\bullet)$, in the notation proposed by Brandwood [1983]) to the Hamiltonian defined here in Equation (2-14), leads to the same eigenvalue/eigenvector problem and to the same solution for the weights presented in Section 3.0. This indicates the generality of the MC formulation, and the equivalency of both gradient operators in the context of the MC approach.

APPENDIX A: ANALYTIC DERIVATION OF THE PERFORMANCE INDEX

The performance index used in the derivation of the MC adaptive sidelobe canceling technique is derived herein using an analytic approach based on the real-valued vector representation of complex-valued variables, which complements the approach adopted in Section 2.0.

The performance index used in the derivation of the MC weights for an adaptive sidelobe canceling problem as presented in Section 2.0 is of the form

$$J(v,\underline{w}) = \frac{1}{2}(\rho_{\alpha\beta} + \overline{\rho}_{\alpha\beta}) = \frac{1}{2}(\underline{w}^{\dagger}\underline{a}v + \underline{w}^{T}\overline{\underline{a}v}) = \mathrm{Re}\{\underline{w}^{\dagger}\underline{a}v\} \tag{A-1}$$

where the complex-valued scalar $\rho_{\alpha\beta}$ is the correlation coefficient to be maximized, $\underline{w}$ and v are the complex-valued weights to be determined, and $\underline{a}$ is the cross-covariance of the signals in the auxiliary channels, $\underline{x}$, and the signal in the main channel, y (see Figure 1-1). More specifically, $\rho_{\alpha\beta}$ is the correlation coefficient of the complex-valued, unity-variance scalars α and β, both of which are linear functions of the channel data,

$$\rho_{\alpha\beta} = E\left[\overline{\alpha}\beta\right] = E\left[\underline{w}^{\dagger}\overline{\underline{x}}yv\right] = \underline{w}^{\dagger}E[\overline{\underline{x}}y]v = \underline{w}^{\dagger}\underline{a}v \tag{A-2}$$

The complex-valued cross-covariance $\underline{a}$ can be represented in terms of its real and imaginary components as

$$\underline{a} = \underline{a}_r + j\,\underline{a}_i = E[\overline{\underline{x}}y] = E\left[(\underline{x}_r - j\,\underline{x}_i)(y_r + j\,y_i)\right] \tag{A-3}$$

$$\underline{a}_r = E[(\underline{x}_r y_r + \underline{x}_i y_i)] \tag{A-4}$$

$$\underline{a}_i = E[(\underline{x}_r y_i - \underline{x}_i y_r)] \tag{A-5}$$

Analogously, the complex-valued scalars α and β have real and imaginary parts that are defined as the following functions of the real and imaginary components of the data and the weights:

(A-6) $$\alpha = \underline{w}^T\underline{x} = \alpha_r + j\,\alpha_i$$

(A-7) $$\alpha_r = \mathrm{Re}\{\underline{w}^T\underline{x}\} = \underline{w}_r^T\underline{x}_r - \underline{w}_i^T\underline{x}_i$$

(A-8) $$\alpha_i = \mathrm{Im}\{\underline{w}^T\underline{x}\} = \underline{w}_i^T\underline{x}_r + \underline{w}_r^T\underline{x}_i$$

(A-9) $$\beta = vy = \beta_r + j\,\beta_i$$

(A-10) $$\beta_r = \mathrm{Re}\{vy\} = v_r y_r - v_i y_i$$

(A-11) $$\beta_i = \mathrm{Im}\{vy\} = v_i y_r - v_r y_i$$

Now define two real-valued, two-element vectors as:

(A-12) $$\underline{\alpha} = \begin{bmatrix} \alpha_r \\ \alpha_i \end{bmatrix}$$

(A-13) $$\underline{\beta} = \begin{bmatrix} \beta_r \\ \beta_i \end{bmatrix}$$

These two vectors constitute a real-valued representation of the complex-valued scalars α and β. Consider now the expected value of the inner product of these two real-valued vectors,

(A-14) $$E[\underline{\alpha}^T\underline{\beta}] = E[\alpha_r\beta_r + \alpha_i\beta_i] = E[\alpha_r\beta_r] + E[\alpha_i\beta_i]$$

Expanding the right-hand-side of Equation (A-14) and applying the expected value operator to products of the random signals $\underline{x}$ and y leads to the following expressions (after substitution of the definitions in Equations (A-4) and (A-5)):

(A-15a) $$E[\underline{\alpha}^T\underline{\beta}] = E\left[\underline{w}_r^T\underline{x}_r y_r v_r - \underline{w}_r^T\underline{x}_r y_i v_i - \underline{w}_i^T\underline{x}_i y_r v_r + \underline{w}_i^T\underline{x}_i y_i v_i\right] + E\left[\underline{w}_i^T\underline{x}_r y_r v_i + \underline{w}_i^T\underline{x}_r y_i v_r + \underline{w}_r^T\underline{x}_i y_r v_i + \underline{w}_r^T\underline{x}_i y_i v_r\right]$$

(A-15b) $$E[\underline{\alpha}^T\underline{\beta}] = \underline{w}_r^T\, E[\underline{x}_r y_r + \underline{x}_i y_i]\, v_r + \underline{w}_i^T\, E[\underline{x}_r y_r + \underline{x}_i y_i]\, v_i + \underline{w}_i^T\, E[\underline{x}_r y_i - \underline{x}_i y_r]\, v_r - \underline{w}_r^T\, E[\underline{x}_r y_i - \underline{x}_i y_r]\, v_i$$

(A-15c) $$E[\underline{\alpha}^T\underline{\beta}] = \underline{w}_r^T\underline{a}_r v_r + \underline{w}_i^T\underline{a}_r v_i + \underline{w}_i^T\underline{a}_i v_r - \underline{w}_r^T\underline{a}_i v_i = \mathrm{Re}\{\underline{w}^\dagger \underline{a} v\}$$

Notice that the right-hand-side of Equation (A-15) is identical to the right-hand-side of Equation (A-1). Therefore, the performance index defined in Section 2.0 is equivalent to a performance index based on the inner product of the real-valued vector representation of the complex-valued scalars α and β. This result provides another justification for the performance index because the expected value of the magnitude of a two-dimensional random vector determines the total variance (power) of the random vector. That is, for a zero-mean, two-dimensional random vector $\underline{s}$,

(A-16) $$\underline{s} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

the total variance is determined as

(A-17) $$E[\,|\underline{s}|^2\,] = E[\underline{s}^T\underline{s}] = E[s_1 s_1 + s_2 s_2] = E[s_1 s_1] + E[s_2 s_2]$$

If $\underline{s}$ represents a vector error quantity, then Equation (A-17) represents the error variance, and minimization of such mean-square error is the objective in a large class of problems.

In summary, the performance index adopted in the derivation of the MC criterion for the adaptive sidelobe canceling problem (Section 2.0) can be interpreted also as the total cross-

covariance of the real-valued, two-dimensional vectors $\underline{\alpha}$ and $\underline{\beta}$ defined in Equations (A-12) and (A-13). That is,

$$J(v,\underline{w}) = \frac{1}{2}\,(\rho_{\alpha\beta} + \overline{\rho}_{\alpha\beta}) = \frac{1}{2}\left(E\left[\overline{\alpha}\beta\right] + E\left[\overline{\beta}\alpha\right]\right) = E\left[\underline{\alpha}^T\underline{\beta}\right] \qquad \text{(A-18)}$$

From this viewpoint, the largest-valued possible maximum of the performance index $J(v,\underline{w})$ in (A-18) occurs when the vectors $\underline{\alpha}$ and $\underline{\beta}$ are co-linear, in which case the complex-valued scalars α and β are linearly proportional.

REFERENCES

T. W. Anderson

(1958) An Introduction to Multivariate Statistical Analysis, John Wiley & Sons, New York, NY.

D. H. Brandwood

(1983) "A complex gradient operator and its application in adaptive array theory," IEE Proceedings, Vol. 130, Parts F and H, No. 1 (February), pp. 11-16.

F. R. Gantmacher

(1977) The Theory of Matrices, Vol. 1, Chelsea Publishing Co., New York, NY.

H. Hotelling

(1936) "Relations between two sets of variates," Biometrika, Vol. 28, pp. 321-377.

J. R. Román

(1991) Adaptive Sidelobe Canceling Using Canonical Variables, Technical Memorandum No. SSC-TM-91-01 (December 30), Scientific Studies Corporation, Palm Beach Gardens, FL.

(1992) On The Performance Index For Adaptive Sidelobe Canceling Using Canonical Variables, Technical Memorandum No. SSC-TM-92-01 (March 27), Scientific Studies Corporation, Palm Beach Gardens, FL.

M. R. Spiegel

(1964) Theory and Problems of Complex Variables, McGraw-Hill, New York, NY.